

Propuesta para una Lingüística del Hipertexto o Ciberlingüística

Por

JOHN ALEXANDER ROBERTO RODRÍGUEZ*
UNIVERSITAT POMPEU FABRA



Gracias a que en la actualidad la idea del lenguaje tal y como era pensado en los comienzos de la lingüística ha evolucionado extraordinariamente para extenderse a otro tipo de manifestaciones menos convencionales, el hipertexto tiene cabida como objeto de la reflexión y de la práctica lingüísticas. El trabajo con hipertextos nos acerca al reto de dotar de una forma adecuada a los contenidos hipertextuales. La navegación, interactividad, escritura, lectura, generación, integración, representación y diseño hipertextual constituyen problemas técnicos abiertos a las prácticas lingüísticas. Todo esto nos hace ver la necesidad de una disciplina que se encargue del estudio y el desarrollo del hipertexto en cuanto unidad lingüística.

Palabras Claves: Ciberlingüística, Lingüística del Hipertexto, Lingüística Aplicada, Lingüística Computacional, Hipertexto.

1. PROPUESTA

Si hablar de Lingüística Computacional (LC), Ingeniería Lingüística e incluso Lingüística Aplicada (LA) puede resultar algo desconcertante para quienes de una manera u otra se dedican al estudio del lenguaje, más desconcertante aún les puede parecer el término Lingüística del Hipertexto (LH). Por este motivo, y aunque mi objetivo es hablarles de ésta última veo provechoso describir en qué consiste cada una de ellas ya que comparten intereses y métodos que es indispensable conocer antes de referirme a la LH.

Para comenzar señalaremos que la **Lingüística Computacional (LC)** es una nueva disciplina académica que surge como consecuencia de los desarrollos

* johnroberto2001@yahoo.es Licenciado en Filología (español e inglés) Master en Lingüística española. Doctorando en Lingüística Aplicada. Instituto Universitario de Lingüística Aplicada (IULA) Universidad Pompeu Fabra (UPF) Línea de investigación: Lingüística Computacional.

• Artículo recibido el 08-mar-04 y aprobado el 18-may-04.

recientes en tecnología de la información y que conecta los conocimientos teóricos en el campo de la lingüística con los métodos y estrategias de la informática. En términos específicos la LC desarrolla modelos formales para simular aspectos de la facultad humana del lenguaje que puedan ser implementados mediante programas de computador. La meta a largo plazo en LC está en desarrollar aplicaciones informáticas inteligentes basadas en la comprensión profunda del lenguaje humano.

Con el objetivo de simular la competencia lingüística y representar los procesos cognitivos que constituyen el uso del lenguaje, la LC trabaja en cooperación con otras disciplinas como la inteligencia artificial (IA), la ciencias cognitivas, la matemática, la lógica, las diferentes filologías, la filosofía o las tecnologías de la información.

Los métodos, técnicas, herramientas y aplicaciones en esta área se localizan bajo un campo más general, el de la **Ingeniería Lingüística** o tecnologías del lenguaje. Las tecnologías del lenguaje dan la posibilidad a la LC de desarrollar software con aplicaciones reales que simplifican el trabajo e incrementan la productividad en diferentes sectores. Algunas de estas aplicaciones consisten en la recuperación de la información, desarrollo de sistemas expertos e interfaces en lenguaje natural, máquinas de traducción automática, bases de datos interrogables en lenguaje natural, sistemas de síntesis de voz, sistemas de respuesta automática mediante voz, editores de texto, robots, etc.

El trabajo en LC puede tener una motivación científica o tecnológica. Es científica cuando provee una explicación computacional para los diferentes fenómenos lingüísticos y psicolingüísticos. Es tecnológica cuando desarrolla sistemas que solucionen problemas específicos cotidianos que involucren el manejo del lenguaje.

A todo esto surge un interrogante, ¿qué debe aprender un estudiante de LC? Como hemos visto la LC no es una material independiente, por lo que a nivel de aprendizaje es necesario adquirir conocimientos tanto en lingüística teórica y empírica como en informática y programación. En la práctica un estudiante que desee dedicarse a resolver problemas en procesamiento del lenguaje natural deberá saber de lógica, programación, estadística, lingüística teórica y lingüística formal, entre otras cosas. Es aquí, justamente, donde se está exigiendo una renovación a nivel de los contenidos universitarios para las filologías y la lingüística. Algunas universidades en Europa pero sobre todo en Estados Unidos llevan va-

rios años promoviendo programas de postgrado e incluso a nivel de pregrado se pueden encontrar iniciativas particulares que incluyen los lenguajes de programación o la lógica entre sus contenidos curriculares.

Dado que tanto la LC como su vertiente más informatizada, la Ingeniería Lingüística, tratan de ofrecer soluciones a problemas lingüísticos, éstas disciplinas conforman el nuevo marco de la **Lingüística Aplicada (LA)**. En sentido general diremos que a la LA atañe el funcionamiento social del lenguaje en situaciones concretas de comunicación, como herramienta para resolver las necesidades comunicativas que la sociedad se plantea. Es, efectivamente, esa tendencia a resolver problemas lo que la caracteriza. Para ser más precisos diremos que la LA tiene como meta fundamental la aplicación de los conocimientos sobre la lengua para resolver una necesidad social.

La etiqueta de “ciencia aplicada” está asociada a la **resolución de problemas reales** a diferencia de la “ciencia teórica” ligada a la **producción de conocimiento**. No obstante, en la práctica científica cabe dejar claro que estas fronteras no son absolutas pues tan cierto es que la resolución de problemas comporta de alguna manera la producción teórica como que ésta es potencialmente objeto de aplicación en la solución de problemas, lo que se suele llamar la “aplicabilidad”.

Dichos presupuestos conducen a ver el criterio de “finalidad científica” como la vía más adecuada para reconocer el carácter teórico o aplicado de una disciplina. Frente a este, el “objeto de estudio” como criterio pierde relevancia, ya que no obliga o compromete en grado alguno el modo de enfrentarse con un tema. La finalidad científica en lingüística, por su parte, bien puede orientarse al conocimiento y a la producción de teoría, o bien, “cifrar sus objetivos en el alcance de soluciones a problemas materiales planteados por las lenguas” (Fernández Pérez, 1996: 13).

Eliasson (1987) plantea claramente esta polaridad científica al reconocer la existencia de la Lingüística aplicada pero, al mismo tiempo, proyectar las dos vertientes a todas las disciplinas lingüísticas: psicolingüística teórica y psicolingüística aplicada, sociolingüística teórica y sociolingüística aplicada, etc. Lo cual nos lleva a afirmar con Milagros Fernández que:

“el terreno de la Lingüística aplicada no se reconoce como un ámbito específico y particular integrado por áreas peculiares y propias, sino que por el contrario aquel terreno se toma como una proyección puntual de las zonas establecidas y deslindadas en la Lingüística teórica.” (Fernández Pérez, M., 1996: 14).

En éste contexto de la nuevas tecnologías del lenguaje se inscribe la siguiente propuesta, recordemos que la evolución de toda disciplina, en particular las más recientes, exigen cambios y adecuaciones. Estos cambios se generan como producto de una necesidad y una necesidad no es otra cosa más que aquello a lo que no podemos sustraernos. Voy a hablarles ahora de un área de conocimiento dispersa, que a pesar de lo descrito en las líneas anteriores, carece de los límites teóricos necesarios para que se desarrolle pero que como estudioso de la lengua y aficionado a las nuevas tecnologías, en particular de Internet, percibo como unidad, como conocimiento y como práctica reales, se trata de la **Lingüística del Hipertexto (LH)** o si se quiere, de una manera más ostentosa, ciberlingüística. Ahora, para entender en qué consiste esta propuesta es indispensable ofrecer una definición del hipertexto y de Internet, su pariente más cercano.

Internet se originó en la palabra inglesa International Network (Red Internacional) con la que se pretendió designar una gran agrupación internacional de redes informáticas que comunicaban las universidades y centros de investigación de todo el mundo. Básicamente, **Internet es un medio de comunicación que usa el lenguaje como vehículo para transmitir información**, por lo tanto es un medio idóneo para el ejercicio de la lingüística que tradicionalmente se ha definido como “the scientific study of language” (Lyons, 1990: 1).

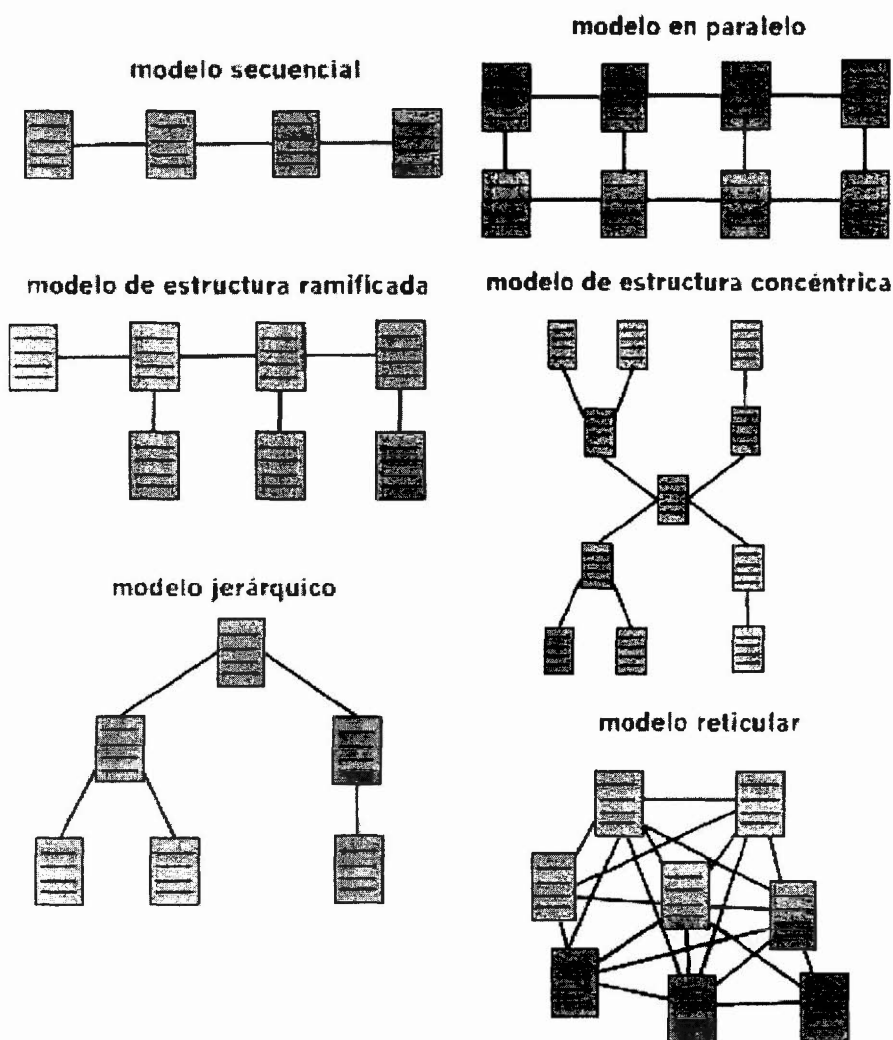
La evolución técnica y conceptual que hoy en día ha logrado Internet lo han convertido en un medio fundamental para la interacción social a todos los niveles. En este medio, el lenguaje ha adquirido connotaciones propias. En efecto, aunque esta nueva forma de comunicación se basa en las formas lingüísticas tradicionales ha desarrollado sus propias peculiaridades, tanto es así que algunos investigadores prefieren hablar de **hiperlenguaje** o **ciberlenguaje** cuando se refieren al lenguaje en Internet, más concretamente al lenguaje del hipertexto, veamos en qué consiste este último.

En Internet la información adopta la forma de una Página Web. Una página web es un documento constituido por información textual, gráfica o sonora integradas mediante un lenguaje de marcado, el HTML o HyperText Markup Language (Lenguaje de Marcas de Hipertextos). El HTML establece las características y posición de los contenidos en las páginas, dichos contenidos adoptan múltiples formas. Si una página web contiene información textual se denomina hipertexto, si contiene información gráfica y sonora se le denomina hipermedia, aunque en la práctica se suele emplear la palabra hipertexto para denominar ambas realidades.

El hipertexto –o hiperdocumento– está compuesto de varios elementos (nodos, contenido, enlaces y propiedades) organizados en una estructura hipertextual y una serie de operaciones. Un **nodo** es un bloque de información que incluye contenidos textuales o gráficos a propósito de un concepto sobre el que el nodo trata. Los nodos son elementos generales que integran los contenidos y las propiedades. El **contenido** hace referencia a la información en sí y puede estar expresada en forma de texto, imágenes o sonido, siendo el texto el componente principal. Las **propiedades** hacen referencia a la forma en como se presentan los contenidos: colores, estructura, tipo de letra, etc. Los **enlaces** son conexiones entre nodos que ofrecen rutas de navegación entre nodos conceptualmente relacionados. Están indicados en la pantalla mediante palabras subrayadas (hot-word), gráficos o iconos (hot-spot).

La combinación de estos elementos puede dar pie a siete **estructuras** básicas para la representación del hipertexto definidas por el tipo de navegación que ofrecen: lineal, ramificada, concéntrica, paralela, jerárquica, reticular y mixta (ver FIGURA 1). En una estructura o modelo hipertextual lineal la única forma posible que existe de ir de un nodo a otro es accediendo al nodo posterior o anterior, esto es de manera secuencial, como en el caso del cine en donde las escenas se suceden una después de otra. La estructura ramificada ofrece la posibilidad de navegar por nodos subordinados fijados a sus nodos padres secuenciales. El modelo en paralelo, por su parte, es posible, además de la navegación secuencial el desplazamiento por nodos del mismo nivel, con lo que resulta útil para mostrar diferentes perspectivas de un tema. En la estructura concéntrica una serie de nodos padre y nodos hijo asociados se organizan en torno a un nodo de “entrada” pero sin permitir la navegación entre nodos del mismo nivel. El modelo jerárquico o “en árbol” parte de un nodo general de “entrada” del que se van desprendiendo otros nodos en diferentes grados de subordinación. Éste es el modelo clásico de organización temática de la información y el empleado por los directorios de la web. El modelo reticular lleva a su máximo nivel de representación el concepto de red pues en él todos los nodos están conectados sin privilegiar un recorrido sobre otro. Finalmente, la estructura mixta combina dos o más modelos recogiendo las ventajas como las desventajas de estos.

Además de ser un objeto estructurado, sobre el hipertexto es posible ejercer una serie de acciones u operaciones. Las **operaciones** pueden ser clasificadas en tres grupos: recuperación, edición y personalización. Cada una de estas operaciones constituye una actividad compleja e independiente. La recuperación consiste

FIGURA 1. Estructuras básicas para la representación del hipertexto.

en el conjunto de tareas que permite localizar y acceder a aquellos hiperdocumentos que son relevantes para solventar una necesidad de información. En la actualidad los sistemas y estrategias para la recuperación han evolucionado sorprendentemente gracias a los avances en el campo de la IR (Information Retrieval) lo cual facilita el trabajo de los usuarios y mejora los resultados. La edición, por su parte, se refiere tanto a las fases de creación como actualización

y eliminación de los hiperdocumentos. Esta etapa comporta la participación de programadores, diseñadores, maquetadores y otros tantos especialistas. La última operación es la personalización y hace referencia al documento final que cada usuario va elaborando en el proceso de búsqueda.

El *hipertexto* es texto inacabado al que se accede de forma no secuencial. Esto quiere decir que se puede “navegar” por el indefinidamente sin seguir un orden lógico –hablando en términos del modelo de lectura que se concibe en la sociedad actualidad.

En cuanto texto, el hipertexto se compone de signos organizados de forma significativa. El hipertexto como cualquier texto consta de palabras que forman oraciones y éstas párrafos y discursos y por lo tanto puede ser objeto del análisis lingüístico. Pero a diferencia del texto (tradicional), el hipertexto se concibe en un ambiente virtual, en el que la tecnología adquiere un papel fundamental. La forma y los usos lingüísticos hipertextuales no son los mismos que los del texto tradicional. En este contexto y gracias a que la idea del lenguaje tal y como era pensado en los comienzos de la lingüística ha evolucionado extraordinariamente para extenderse a otro tipo de manifestaciones menos convencionales, el hipertexto tiene cabida como objeto de la reflexión y de la práctica lingüísticas. A la lingüística en general, y a una rama de la lingüística en particular, concierne, el estudio tanto formal como funcional del *hipertexto*.

Luego de pasar revisión de los conceptos relevantes e ideas acerca de la naturaleza comunicativa de la Internet y en concreto del hipertexto, es hora de aportar una definición de lo que en términos didácticos hemos dado en denominar Lingüística del Hipertexto.

En términos generales **definiremos la Lingüística del Hipertexto como un conjunto de prácticas tendientes a la comprensión y desarrollo lingüístico del hipertexto**. No se pretende, con ello, concederle el estatus de disciplina autónoma aunque sí reconocer que tiene un campo propio en cuanto constituye una práctica ligada a la solución de necesidades sociales y trabaja sobre un objeto: el hipertexto. Esta caracterización no niega que deba haber teoría en el ámbito de la Lingüística del Hipertexto, pero siempre y cuando esté enfocada a resolver los problemas materiales que plantee el hipertexto.

Una definición de este tipo nos conduce a reflexionar sobre la naturaleza lingüística del hipertexto. Los trabajos existentes en este sentido tienden a ca-

racterizar al hipertexto a través de sus propiedades discursivas. Esto se debe a la consciencia que se tiene del hipertexto como una **unidad enunciativa de la que es imposible, en la práctica cotidiana, interpretar por separado imágenes, texto y sonido**: “El pensamiento humano no es sólo verbal ni puede ser únicamente reducido a texto” (Noci, 2001: 90). Los espacios digitales, aseguran los estudiosos del hipertexto, imponen otros hábitos de escritura y lectura, igualmente proponen nuevas formas de retórica, de lenguaje y de discurso. Como dice Gabriela Häbich, “si los espacios digitales han de permitir una nueva forma escritural, ésta probablemente pasará por un espacio ideográfico dinámico en el que convivan logogramas, pictogramas, escrituras fonéticas y alfabéticas.” y Jean Cloutier, “se abre la posibilidad de un lenguaje ‘polisintético e integrador’. Polisintético porque aglutina diversos elementos, que toman sentido sólo dentro del conjunto.” (Noci, 2001: 87 y 90). Ante tales evidencias, doptaremos las denominaciones “**discurso hipertextual**” (Noci, 2001: 107; Clément) e “**hiperlenguaje**” (Fernández Serrano) para designar ese par de realidades que conciernen a la Lingüística del Hipertexto.

De esta manera el hiperlenguaje a nivel textual puede ser abordado por la lexicografía, la gramática, la semántica o la pragmática; a nivel gráfico por la semiótica; y a nivel sonoro por la fonética y fonología.

De manera específica, a la **psicolingüística** competiría, entre otras cosas, el análisis de las estrategias de comprensión y producción del hipertexto; a la **sociolingüística**, explicar las relaciones hiperlenguaje-sociedad; a la **pragmática**, el uso intensivo del lenguaje en Internet; a la **semiótica**, el análisis de la “gramática de las imágenes” en el hiperdocumento, a los analistas del **discurso** el quién o quienes utilizan el hiperlenguaje, cómo lo usan y por qué los usan, etc. Todas estas actividades estarían encaminadas, repetimos, a contribuir al desarrollo eficiente del hipertexto en cuanto unidad lingüística y discursiva.

Las aproximaciones lingüísticas actuales al hipertexto se han dado mayoritariamente en el terreno formal. En esta línea son importantes los trabajos en la llamada “gramática hipertextual” (hypertextual grammar) y la “retórica del hipertexto” (rhetoric of hypertext). La primera busca facilitar la generación automática de vínculos considerando su forma y contenido semántico; la segunda, trabaja sobre la estructura hipertextual inmiscuyendo tanto a su autor como a su lector. Una serie de aproximaciones emparentadas con éstas son: estructura del hipertexto (Hypertext structure), hipertexto adaptativo (Adaptive Hypertext), hipertexto y cognición (Hypertext and knowledge), linealidad (Linearity), Web

semántica (Semantic Web) y Recuperación de la información (Information Retrieval) (Balpe, Jean-Pierre, 1996).

A nivel teórico prevalecen sobre todo los estudios discursivos que o bien reinterpretan los conocimientos existentes sobre el discurso para aplicarlos al hipertexto, o bien, aplican dichos conocimientos sin reinterpretación alguna. Así, George Landown sostiene que:

el hipertexto nos ofrece un excelente campo de ensayo para poner a prueba nuestras teorías sobre el discurso, sobre la pragmática lingüística y el estatuto epistemológico de los actos de leer y escribir" (Fernández Serrano: 2).

Aunque para Landown las implicaciones discursivas giran en torno al uso social del hipertexto dado que presupone la existencia de discursos previos vinculados cada uno con su propia semiótica, lo fundamental al parecer de otros es su naturaleza discursiva, para ello se basan en las conclusiones de los teóricos de la filosofía y el lenguaje como Michel Foucault, Paul Ricoeur, Roland Barthes o Jacques Derrida quienes nos ha ayudado a entender el texto:

En efecto, ya sabemos que un texto es una estructura dinámica y que desde un punto de vista pragmático, en el momento de su actualización interpretativa por parte de un destinatario las marcas de comienzo y fin no son fronteras de cláusula, sino puntos de conexión con otros textos que el destinatario puede poner o no en funcionamiento. Sabemos que todo texto nace contextualizado psíquica, cultural, política, socialmente. Sabemos que la polifonía y la intertextualidad como estrategia que van tejiendo el discurso, sea cual sea su cristalización en texto. Sabemos que la lectura es un acto de cooperación semiótica y que los roles de autor/lector son complementarios insolubles. Sabemos de experiencias de textos abiertos, de narrativas discontinuas y cronológicamente no lineales, de discursos "rizomáticos", y hasta "autodestructivos", sabemos de la interactividad efectiva autor/lector, de coautoría escritural, o del diálogo de varios códigos y semióticas en un mismo texto, etc. (Fernández Serrano: 4)

Algunos autores son más tajantes: "Lo cierto es que los sistemas y entornos informáticos hipertextuales e hipermedia no añaden nada nuevo a las actuales teorías del discurso..." (Fernández Serrano: 5). Otros, por su parte, no comparten esta idea: la Lingüística, la Semántica, el Análisis del discurso y Análisis de contenido aportan propuestas de esquemas de comprensión y producción de documentos, y, por ende, modos objetivados de organizar el conocimiento. (García Jiménez: 2).

No obstante para quienes trabajamos en el desarrollo técnico del hipertexto, siempre con un enfoque lingüístico, evidenciamos la necesidad de integrar estos esfuerzos algo dispersos en un marco académico en el que se puedan compartir conocimientos pertinentes a nuestros intereses y necesidades. Recogiendo lo dicho hasta el momento tenemos que dentro de las tareas de la LH estarían:

- análisis y producción de nuevas formas de retórica, de lenguaje y de discurso
- análisis de las estrategias de comprensión y producción del hipertexto
- explicación de las relaciones hiperlenguaje-sociedad
- uso intensivo del lenguaje en Internet: quién utiliza el hiperlenguaje, cómo lo usa y por qué lo usa
- análisis de la gramática de las imágenes en el hipertexto
- generación automática de vínculos y contenidos
- análisis de la estructura hipertextual
- diseño de interfaces ergonómicas e inteligentes

2. CONCLUSIÓN

Las nuevas formas que asume el lenguaje en ambientes virtuales (hipertexto) lleva a que los lingüistas se interesen cada vez más por su estudio y permite a los lingüistas aplicados proponer estrategias adecuadas para su recepción y producción. Al reconducir estas prácticas en términos de necesidades nos vemos inmerso en un campo inexplorado de la lingüística que hemos denominado Lingüística del Hipertexto y concebido como un conjunto de prácticas tendientes a la comprensión y desarrollo lingüístico del hipertexto.

Lo expuesto en este artículo ha de ser asumido como una propuesta abierta a los investigadores, sobre todo lingüistas, que puedan estar interesados en contribuir al desarrollo de los principios epistemológicos de ésta disciplina. Todos los trabajos en el campo de la lingüística y el hipertexto que no han encontrado hasta la fecha un espacio óptimo para desplegar sus ideas, con el tiempo tendrán cabida en una Lingüística del Hipertexto (LH) o Ciberlingüística (CL).

REFERENCIAS

- BALPE, Jean-Pierre *et al.* (1996). *Techniques avancées pour l'hypertexte*. París: Hermes. 288 p.
- FERNÁNDEZ PÉREZ, Milagros *et al.* (1996). *Avances en lingüística aplicada*. [coordinadora Milagros Fernández Pérez]. Santiago de Compostela: Universidad Santiago de Compostela. 372 p.
- GARCÍA JIMÉNEZ, Antonio. (s. d.). *La gestión y el tratamiento de la información hipermedia en las instituciones: organización del conocimiento e información electrónica*. Universidad

Europea de Madrid: VI Jornadas Españolas de Documentación. http://fesabid98.florida-uni.es/Comunicaciones/A_Garcia.htm

LANDOW, George. (1995). **Hipertexto. La convergencia de la teoría crítica contemporánea y la tecnología**. Paidós: Barcelona. 288 p.

LYONS, John. (1990). **Language and linguistics. An introduction**. Great Britain: Cambridge University Press.

NOCI DÍAZ, Javier. (2001). **La escritura digital. Hipertexto y construcción del discurso informativo en el periodismo electrónico**. Gipuzkoa: Universidad del País Vasco. 233 p.

SERRANO FERNÁNDEZ, Juan. (s. d.). **Hipertexto electrónico e hiperlenguaje, efectos discursivos y refracciones ideológicas**. GITTCUS. <http://www.uned.es/ntedu/espanol/master/primero/modulos/teoria-de-la-informacion>.